# 基于结构熵的警示传播算法收敛性分析 *

牛　进，王晓峰，林青文

(北方民族大学 计算机科学与工程学院，银川 750021)

**摘　要：**收敛性是评价信息传播算法性能的重要指标，信息传播算法求解可满足性问题时，命题公式的结构特征影响算法的收敛性。具有复杂结构的命题公式，信息传播算法不总收敛，为了系统地对此现象给予理论解释。借助于结构熵的方法和技术，提出命题公式的结构熵模型及其度量方法，计算随机可满足性实例的结构熵。警示传播算法(WP)作为信息传播算法的基本模型，分析 WP 算法的收敛性对于研究其他信息传播算法的收敛性具有重要意义，分析了 WP 算法收敛性与结构熵之间的关系，给出 WP 算法收敛的判定条件。通过实验分析，该方法有效可行。

**关键词：**可满足性问题；命题公式；结构熵；警示传播算法；收敛性；

**中图分类号：**TP　　doi: 10.19734/j.issn.1001-3695.2020.03.0053

# Convergence analysis of warning propagation algorithm based on structural entropy

Niu Jin, Wang Xiaofeng, Lin Qingwen

(*School of Computer Science & Engineering*, *North Minzu University*, *Yinchuan* 750021, *China*)

**Abstract:** Convergence is an important index for evaluating the performance of information propagation algorithms. When the information propagation algorithm solves the satisfiability problem, the structural characteristics of the proposition formula affect the convergence of the algorithm, propositional formulas with complex structures do not always converge on information propagation algorithms. For this phenomenon, there are few systematic theoretical explanations. As the basic model of the information propagation algorithm, the warning propagation algorithm (WP) , analyzing the convergence of the WP algorithm is of great significance to study the convergence of other information propagation algorithms. With the help of structural entropy methods and techniques, people proposed a structural entropy model of propositional formula and its measurement method. This paper calculated the structural entropy of a random satisfiability instance, analyze the relationship between the convergence of the WP algorithm and the structural entropy, and give the judgment condition of convergence of WP algorithm. Through experimental analysis, this method is effective and feasible.

**Key words:** satisfiability problem; propositional formula; structural entropy; warning propagation algorithm; convergence;

## 0　引言

约束可满足性问题(constraint satisfaction problem)普遍存在于现实世界中，其研究目前在人工智能领域中具有广泛影响[1~3]，作为一个重大子问题，布尔可满足性判定问题(satisfiability,SAT)作为首先被证明的 NPC 问题[4]，任何算法都无法在多项式时间内求得准确解，然而，现实世界中的许多问题都可以规约为 SAT 问题[5~7]。因此，SAT 问题一直在计算机科学界和数理逻辑界被学者广泛关注。该问题研究是否存在一组布尔变量，赋值给指定的命题公式，能使该公式可满足，即使公式为真。由于任何一个命题逻辑公式都可以转换为一个等价的合取范式[8]，因此本文只考虑合取范式。

20 世纪 80 年代，物理学家提出了一种基于因子图的信息传播算法，该算法的性能十分依赖于因子图的结构，当因子图为单连通结构(任意两个节点之间只有一条路径)时，算法一般能够正确收敛。虽然信息传播算法用于求解 SAT 问题时往往具有良好的性能表现，但随着因子图结构趋于复杂，尤其当其中出现环路时，信息可能在环路中无限循环传播，该过程将导致信息最终无法收敛于一个固定值，因此导致算法无法求解，该现象称为算法不收敛[9]。WP 算法作为一种最为基础的信息传播算法，研究其收敛性可以对其他信息传播算法的研究给出重要参考。WP 算法是一种基于因子图的消息迭代算法，通过因子图的边进行信息传递(称为警示信息)，警示信息表示变元的取值是否能够完全决定子句的可满足性，该算法利用一种迭代策略更新警示信息，直到算法收敛时停止迭代，或达到最大迭代步数时若算法仍不收敛，则强迫迭代过程结束，因此收敛性是 WP 算法最为核心的性质。目前，WP 算法的收敛性研究已经获得了一些成果，文献[10]证明了当因子图中只存在一个环路时，该算法能够有效收敛,得出的值是变量边缘分布的有效近似。文献[11]基于高斯模型给出了算法收敛的充分必要条件,该条件受限于两个变量不能同属于一个子句。文献[12]基于消息矩阵分析了算法的收敛性,给出了一个充分必要条件,但局限于消息矩阵为半正定矩阵。文献[13,14]分析了基于高斯模型的算法的收敛性,但要求用半定规划来检查收敛性,检查效率较低且非常困难。文献[15]分析了求解均值的算法的收敛性,要求估计有限尺度矩阵的谱半径,在实际应用中可行性不高。文献[16]中分析了算法的收敛性，并给出了信息传播算法收敛的一个概率条件，但该条件主要局限于植入指派的随机可满足实例产生模型 $P_{n,p}^{plant}$ 。在 $P_{n,p}^{plant}$ 模型中要求概率 p 充分大，其本质是高概率地最多只有一个指派满足该实例。不难看出,上述关于警示传播算法的收敛性分析都是基于某个具体模型,或者使用范围窄,或者判

---

**收稿日期：**2020-03-25；**修回日期：**2020-05-23　**基金项目：**国家自然科学基金资助项目(61462001，61762019，61862051，61962002)；北方民族大学重大专项(ZDZX201901)；宁夏自然科学基金资助项目(NZ17111，2019AAC03120，2019AAC03119)；北方民族大学校级科研一般项目(2019XYZJK05)

**作者简介：**牛进(1993-)，男，陕西安康人，硕士研究生，主要研究方向为机器学习、算法分析与设计.(20187258@stu.nmu.edu.cn)；王晓峰(1980-)，男，甘肃会宁人，副教授，博士，主要研究方向为机器学习、算法分析与设计.(xfwang@nun.edu.cn)；林青文(1996-)，女，黑龙江牡丹江人，硕士研究生，主要研究方向为机器学习、算法分析与设计.

定条件的验证复杂等。因此,对警示传播算法的收敛性分析仍然不完善，尤其基于普通因子图结构的算法收敛性研究尚为缺失。

由于 WP 算法的收敛性十分依赖于因子图的结构，而随着因子图规模的增大，因子图上的节点之间相互传递消息的过程不仅变得极为复杂，同时具有随机不确定性。Shannon[17]提出的信息熵是关于不确定性的一种度量，作为关于概率分布的静态度量方法，在信息论中，若根据概率分布选择节点，则熵值反映了确定该节点的代码所需的平均信息量。所以，在通信领域中，信息熵仅用于度量点与点之间的单一信息。在此基础上，文献[18]首次提出了结构熵。与信息熵不同的是，结构熵是一种动态度量方法，可以反映网络的动态复杂性。因此，结构熵弥补了信息熵无法确切反映网络结构复杂性的缺陷，可以用以度量节点的交互复杂度和内部结构复杂度等重要信息。引入结构熵不仅可以度量 WP 算法在因子图上的复杂性和不确定性等重要结构信息，还可以反映 WP 算法中随机游走过程的动态复杂性。

综上所述，为了研究 WP 算法求解 SAT 问题的收敛性，本文基于结构熵的概念，提出了命题公式的结构信息。利用 $G=(n,3,m)$ 模型随机生成不同规模的 SAT 实例，获取实例的结构信息。结合 WP 算法在实例上的收敛情况，分析 WP 算法的收敛性和结构熵之间的关系，进而利用结构熵度量 WP 算法的收敛性，给出基于结构熵的 WP 算法的收敛条件。

## 1　基础知识

对于一组给定的布尔变量 $V=\{x_1,x_2,\cdots,x_n\}$ ,集合中的每一个元素称为变元，若变元 $x_i$ 在子句中以自身值出现，对应的文字为 $x_i$ (正文字)。若 $x_i$ 在子句中取反出现，对应的文字为 $\neg x_i$ (负文字)。文字的析取可以构成子句，记作 $C_i=(x_1\vee x_2\vee\cdots\vee x_n)$，每个子句中包含的所有非重复文字个数称为子句的长度，记作 $|C|$ [19]。而子句的合取构成合取范式公式(CNF)，记作 $F=(C_1\wedge C_2\wedge\cdots\wedge C_m)$，简记为 $F=(C_1,C_2,\cdots,C_m)$。SAT 问题研究是否存在一组对变元的赋值指派，使每一个子句都为真，即使 CNF 公式为真。特别地，当一个 CNF 公式中任意子句的长度都为 $r$ ,则该 CNF 公式称为 $r-$CNF。由于任何一个 CNF 公式都可以在多项式时间内规约为一个 3-CNF 公式[20]，所以本文只针对 3-CNF 公式进行研究。

### 1.1　因子图

设 CNF 公式为：$F=(C_1,C_2,\cdots,C_m)$ , 变元集合记为 $\{x_1,x_2,\cdots,x_n\}$，用 $i$ 表示 $x_i$。公式 $F$ 可以构造一个二部图 $G=(C\cup X,E)$，称 $G$ 为因子图。其中，$X=\{1,2,\cdots,n\}$ 称为变元集合, $C=\{C_1,C_2,\cdots C_m\}$ 称为子句集合。

因子图中有两类边：

实边: $(C_i,j)\in E\Leftrightarrow$ 子句 $C_i$ 含正文字 $x_i$；

虚边: $(C_i,j)\in E\Leftrightarrow$ 子句 $C_i$ 含负文字 $\neg x_i$。

例如：

$$F=(x_1\vee\neg x_2\vee\neg x_3)\wedge(x_1\vee\neg x_2\vee x_4)\wedge$$
$$(\neg x_2\vee x_3\vee x_5)\wedge(\neg x_2\vee x_4\vee x_5)=a_1,a_2,a_3,a_4$$

的因子图 $G=(V,E)$ 如图 1 所示。

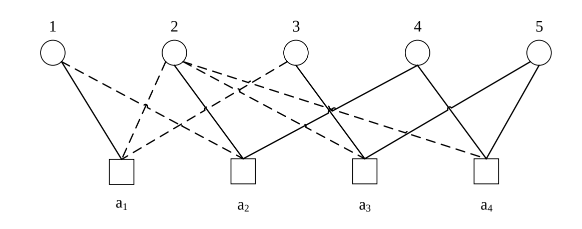



图 1　因子图

Fig. 1　Factor diagram

### 1.2　WP 算法

在因子图的每条边 $(a,i)$ 上，定义 *WP* 算法中的消息 $u_{a\to i}$ (称为警示信息)为从子句节点 $a$ 向变元节点 $i$ 传递的一个布尔值。$u_{a\to i}$ 可以表明变元 $x_i$ 的取值情况能否完全决定子句 $a$ 的可满足性，$u_{a\to i}=1$ 表示：变元 $x_i$ 的取值情况可以决定子句 $a$ 的可满足性；$u_{a\to i}=0$ 表示：变元 $x_i$ 的取值情况对子句 $a$ 的可满足性无法起到决定性作用，即其他变元的取值也能够令子句 $a$ 可满足。*WP* 算法的消息更新迭代方程如下：

$$u_{a\to i}(t)=\prod_{j\in V(a)\backslash i}\theta\left(-J_j^a\left(\sum_{b\in V(j)\backslash a}J_j^b u_{b\to j}(t-1)\right)\right)\tag{1}$$

一般改写为如下形式：

$$u_{a\to i}(t)=\prod_{j\in V(a)\backslash i}\theta\left(-J_j^a\left(\sum_{b\in V^+(j)\backslash a}u_{b\to j}(t-1)-\sum_{b\in V^-(j)\backslash a}u_{b\to j}(t-1)\right)\right)\tag{2}$$

$$h_{j\to a}=\sum_{b\in V^+(j)\backslash a}u_{b\to j}(t-1)-\sum_{b\in V^-(j)\backslash a}u_{b\to j}(t-1)\tag{3}$$

其中，$h_{j\to a}$ 为腔域，若变元 $x_j$ 只在子句 $a$ 中出现，则令 $h_{j\to a}=0$。求解 CNF 公式 $F$ 的 WP 算法如下：

算法 1　警示传播算法

输入：CNF 公式 F 因子图 $G=(V,E)$，最大迭代步 tmax

输出：if　tmax 步后 WP 算法未收敛，

　　输出 UN-CONVERGED

　else:

　　输出所有“警示信息” $u^*_{a\to i}$

begin

a) 在初始时刻 t=0，对因子图 $G=(V,E)$ 中每一条边，分别以 1/2 的概率随机地从{0，1}中取一个值赋给 $u_{a\to i}(0)$；

b) for t=1 to tmax

(a)以某种顺序(如：字典序)，对图 $G$ 中的每一条边调用式(1)更新 $u_{a\to i}$；

(b)if　$u_{a\to i}(t)=u_{a\to i}(t-1)$：

$u^*_{a\to i}=u_{a\to i}(t)$；

输出 $u^*_{a\to i}$ .

endfor

c)输出 UN-CONVERGED；

end；

### 1.3　结构熵

结构熵可以用来度量 WP 算法在因子图上的动态复杂性。因子图上的消息传播情况复杂，具有随机性，并且由于因子图中子句节点与变量节点之间可以通过边直接相连，而子句节点相互之间只能通过其他边间接连接，同样，变量节点之间也只能间接相连。节点之间这样的特殊关系令因子图产生了多层次结构，因此，本文引入结构熵度量此类多层次因子图结构。对给定无向连通图 $G=(V,E)$，设 $\mathrm{T}$ 是 $G$ 的分割树，若 $\alpha\in\mathrm{T}$，且 $\alpha\neq\lambda$，结构熵定义如下：

$$\mathrm{H}^{\mathrm{T}}(G)=\sum_{\alpha\in\mathrm{T},\alpha\neq\lambda}H^{\mathrm{T}}(G;\alpha)=\sum_{\alpha\in\mathrm{T},\alpha\neq\lambda}-\frac{g_\alpha}{2m}\log_2\frac{V_\alpha}{V_{\alpha^-}}\tag{4}$$

其中 $g_\alpha$ 是集合 $T_\alpha$ 内部的节点与该集合之外的节点相连的边数，$V_\beta$ 是集合 $T_\beta$ 的体积，即 $T_\beta$ 中所有节点的度数之和。(对于一个带权图 $G=(V,E)$ )，$g_\alpha$ 是集合 $T_\alpha$ 内部的节点与该集合之外的节点之间的所有边的权重之和。图 $G$ 中节点 $v\in V$ 的度数是入射到 $v$ 的所有边的权重之和。

带权无向连通图 $G$ 的 $K$ 维结构熵定义如下：

$$\mathrm{H}^K(G)=\min_T\{\mathrm{H}^{\mathrm{T}}(G)\}\tag{5}$$

特别地，带权无向非连通图 $G$ 的 $K$ 维结构熵定义如下：

$$\mathrm{H}^K(G)=\frac{1}{Vol(G)}\cdot\sum_{j=1}^{L}Vol(G_j)\cdot\mathrm{H}^K(G_j)\tag{6}$$

其中，$Vol(G_j)$ 是 $G_j$ 的体积，而 $G_1$，$G_2$，⋯，$G_L$ 是 $G$ 的所有连通分量。

## 2　命题公式的结构熵

根据 WP 算法收敛性分析可知，当因子图不是树型结构，而是其中包含多个环路的一般图状结构时，算法的收敛性无法保证，因此本文利用高维结构熵度量因子图的结构复杂性，从而分析研究 WP 算法在含有多环路的一般结构因子图上的收敛性表现。

### 2.1　无向图转换技术

为了计算命题公式的结构熵，需要把命题公式的因子图模型转换为带权值的无向图模型。因子图利用两类节点映射 CNF 公式中的子句和变量，并用虚边和实边表示节点之间的关系，这样的结构可以最直观表达因子图和对应公式的映射以及 WP 算法的运算过程。构建邻接矩阵，将因子图中的两类节点映射为无向图中的一类节点，以权值代替因子图中的节点关系表达方式，具体地讲：

设变量节点为 a 和 b，子句节点为 i 和 j，根据式(1)，不妨令 $u_{a\to i}(t)=1$，则 $\forall j\in V(a)\backslash i$，$\theta\left(-J_j^a\left(\sum_{b\in V(j)\backslash a}J_j^b u_{b\to j}\right)\right)=1$ 即 $J_j^a\left(\sum_{b\in V(j)\backslash a}J_j^b u_{b\to j}\right)<0$，于是存在两种情况，①：$J_j^a=-1且J_j^b=1$，②：$J_j^a=1且J_j^b=-1$。上述两种情况表明，变量 j 在子句 a 和 b 中正负出现的情况相反，根据子句的互异性，在情形①中，存在(1) $J_i^a=1且J_i^b=1$ 或(2) $J_i^a=-1且J_i^b=-1$ 或者(3) $J_i^a=1且J_i^b=-1$。在情形②中，存在(1) $J_i^a=1且J_i^b=1$ 或(2) $J_i^a=-1且J_i^b=-1$ 或者(3) $J_i^a=-1且J_i^b=1$。上述分析表明，$u_{a\to i}(t)=1$ 时，环路中正负文字的比率 $\beta$ 存在两组情形，每组比率都分别为 0.25，0.25 以及 0.5。若 $u_{a\to i}(t)=0$，同理可证，当 $J_j^a\left(\sum_{b\in V(j)\backslash a}J_j^b u_{b\to j}\right)>0$ 时，存在两种情况，①：$J_j^a=-1且J_j^b=-1$，②：$J_j^a=1且J_j^b=1$。根据上述分析，$\beta$ 存在两组情形，每组比率都分别为 0.25 以及 0.25。而当 $J_j^a\left(\sum_{b\in V(j)\backslash a}J_j^b u_{b\to j}\right)=0$ 时，由于结果只受 $u_{b\to j}$ 影响，因此 $\beta$ 包含以上所有可能性，且对最终结果不会产生影响，所以不对这种情况做具体分析。

显然，在基本环路模型中，正文字与负文字的比率期望值为 $E(x)=0.3$，将该期望值设为负文字或正文字的权值。由于结构熵只关心因子图的结构信息，无论虚边权值为 0.3、实边权值为 0.7，还是采取相反赋值，两种赋值产生的图仅为对称结构，不影响结构信息。因此，本文为了统一结果，均设虚边权值为 0.3，实边权值为 0.7。采用此方法，将图 1 因子图转换为带权无向图如图 2 所示。

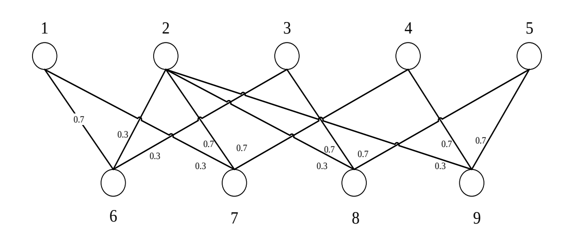



图 2　带权无向图

Fig. 2　Weighted undirected graph

### 2.2　结构熵模型及其度量方法

给定由 $n$ 个节点和 $m$ 个边组成的加权无向连通图 $G=(V,E)$，为了计算结构熵，利用社区划分构造由图中节点集合作为树节点的分割树，该树型结构可以根据以下规定进行构造，其中，$P=\{X_1,X_2,\ldots X_L\}$ 是节点集合 $V$ 的分割，$X_j$ 称为社区：

a)将根节点表示为 $\lambda$，定义节点集合 $T_\lambda=V$；

b)对于每一个节点 $\alpha\in T$，$\alpha$ 的直接后继节点为 $\alpha^{\wedge}j$，$j=\{1,2,\ldots n\}$，节点按照 $j$ 的递增顺序自左向右排序。因此，当且仅当 $i<j$，将 $\alpha^{\wedge}i$ 为 $\alpha^{\wedge}j$ 的左节点记为 $\alpha^{\wedge}i<L\alpha^{\wedge}j$；

c)对每一个节点 $\alpha\in T$，集合 $T_\alpha\subset V$。对于节点 $\alpha$ 和 $\beta$，$\alpha\subset\beta$ 表示 $\alpha$ 是 $\beta$ 的初始字符段。当节点 $\alpha\neq\lambda$，用 $\alpha^-$ 表示 $\alpha$ 的最长初始字符段；

d)对于每个 $i$，$\{T_\alpha\mid h(\alpha)=i\}$ 是 $V$ 的一种分割情况，其中 $h(\alpha)$ 是 $\alpha$ 的高度(根节点 $\lambda$ 的高度为 0，并且对于每个节点 $\alpha\neq\lambda$，$h(\alpha)=h(\alpha^-)+1$；

e)对于每个 $\alpha$，当 $\beta^-=\alpha$ 时，$T_\alpha$ 是 $T_\beta$ 的并集，因此，$T_\alpha=\cup_{\beta^-}=\alpha T_\beta$；

f)对于 T 的每个叶节点 $\alpha$，$T_\alpha$ 是一个单例节点，因此，$T_\alpha$ 包含 $V$ 的单个节点。

初始化分割树后，为计算结构熵，对带权无向图进行最小割分割可以得到两个非重叠社区 $M_1$ 和 $M_2$，将这两个社区作为分割树的第二层树节点，其父节点都为根节点 $\lambda$ (包含因子图中的所有节点，即所有节点划分为一个社区)，其中，每一个社区可能包含一个或多个节点，且 $V=M_1\cup M_2$。若 $M_1$ 或 $M_2$ 包含所节点的数目大于 1，则继续利用最小割对 $M_1$ 或 $M_2$ 进行分割，将分割可能得到的子社区 $M_{11}$，$M_{12}$ 以及 $M_{21}$，$M_{22}$ 作为分割树的第三层树节点，其中，$M_{11}$ 和 $M_{12}$ 为 $M_1$ 的直接后继节点，$M_{21}$ 和 $M_{22}$ 为 $M_2$ 的直接后继节点，且 $M_1=M_{11}\cup M_{12}$，$M_2=M_{21}\cup M_{22}$。若子社区中仍包含一个或多个节点，则重复如上步骤继续分割子社区，并构造分割树，直到当社区之中只包含单个节点时，对该社区停止分割。

若因子图为带权无向非连通图，则遍历其节点，找到每一个连通分量，对其每个连通分量进行如上的图分割并且构造分割树。因子图的最小割过程与对应分割树的构造过程如图 3 和 4 所示(为方便表示，仍然用虚线和实线替代权值)：

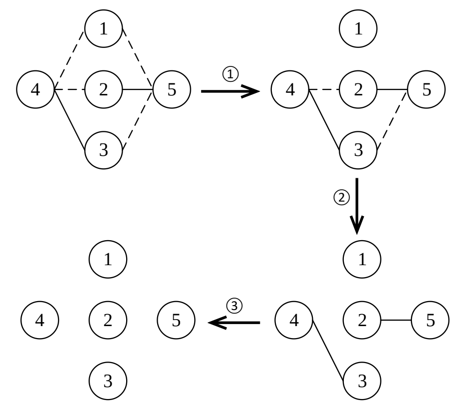



图 3　因子图分割

Fig. 3　Factor graph segmentation

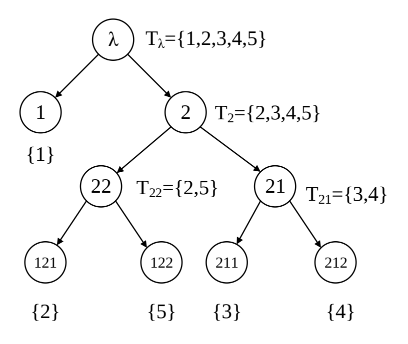



图 4　分割树构造

Fig. 4　Partition tree structure

分割树构造完成后，利用式(4)计算分割树中的每个节点的熵值，其中节点的度数为对应带权无向图中相应节点的度数，对所有节点的熵值作和得到 $H^T(G)$。由式(5)可知，结构熵应取 $H^T(G)$ 的最小值，即找到一种最优分割。目前学术界

对于社区发现的研究正在进行，由于利用最小割算法进行社区发现，社区之间的边密度较小，社区内部的边密度相对较大，所以最小割算法在现阶段社区发现算法当中应用较为广泛[21]。利用在前节中求得正负文字比率的期望，根据该期望值对边赋予权值，因此，基于该权值的最小割不仅能够很好满足结构熵中的社区定义，并且能够最小化结构熵计算过程中的结构信息损耗。因此，上述方法所求结果为 $\min_T\{H^T(G)\}$，即为 CNF 公式的结构信息。

算法 2　CNF 公式结构信息获取

输入：CNF 公式 F

输出：$H^K(G)$

begin

a) $X \leftarrow \{1,2,\cdots,n\}$，$C \leftarrow \{C_1,C_2,\cdots C_m\}$，构造因子图 $G=(C \cup X,E)$；

b) 实线 FL=0.7，虚线 DL=0.3，构造带权无向图 $G(F)$；

c) 初始化分割树，$T_\lambda=\varnothing$；

d) if $G(F)$ connect:

　　go to step e;

　　flag = 1

　else:

　　遍历所有连通分量，go to step e;

　　flag = 0

e) 利用最小割分割图；

f) if Subgraph == single node:

　　go to step g;

　else:

　　go to step e;

g) 利用最小割结果构造分割树；

h) if flag = 1:

$$H^K(G)=\min_T\{H^T(G)\};$$

　else:

$$H^K(G_j)=\sum_{\alpha\in T,\alpha\neq\lambda}H^T(G;\alpha)$$

$$H^K(G)=\frac{1}{Vol(G)}\cdot\sum_{j=1}^{L}Vol(G_j)\cdot H^K(G_j)$$

i) return $H^K(G)$;

end

由式(4)可知，基于分割树的结构熵为分割树中每个节点熵值相加所得，即将不同社区的结构熵求和。本文得到的因子图的结构熵考虑了 WP 算法在因子图上的动态信息随机传播过程，并且，基于该过程所产生的社区结构，利用式(5)可以有效度量社区之间的动态信息交互和社区内部的信息交互过程，而子社区的分割过程基本可以考虑到因子图的所有社区分割情况，从根本上分析了因子图的结构复杂性。结构熵可以有效并完整地分析基于上述图结构的动态复杂性和不确定性等重要性质。利用正负文字比率的期望值作为权值进行最小割，可以有效去除图中的噪声，同时最小化结构信息抽象为结构熵的过程中的信息损耗，该算法求得 CNF 公式结构熵，其中包含着 CNF 公式的结构复杂性，不确定性以及交互性等重要结构信息，最终的计算结果是 CNF 公式的结构信息的高度抽象，并以结构熵的形式表现。

将该算法用于分析 WP 算法收敛性研究中，不仅不用基于特定结构的图模型，使用范围宽泛，判断条件的验证简洁，并且可以填补以往 WP 算法收敛性分析的空白，对基于复杂因子图结构的 WP 算法收敛性研究给予系统的理论支持，为 WP 算法研究提供新的思路与方向。接下来本文将基于结构熵以实验验证分析 WP 算法收敛性。

## 3　WP 算法的收敛性分析

由于比值 $\alpha=m/n$ 对因子图的结构和 WP 算法收敛性具有重要影响，所以本实验利用模型 $G=(n,3,m)$ 产生随机实例。其中，$n$ 代表变元集合中所有变元的个数，$m$ 代表子句个数，$k=3$ 代表每个子句都恰好包含 3 个变元。实验选取两组数据，变元个数分别设置为 $n=10$ 和 $n=20$，令 WP 算法最大迭代次数为 100。以 $n=10$ 的情况为例，具体实验过程为：令 $k=3$，通过固定 $n$ 的规模，利用模型控制 $\alpha$，用以产生不同规模的随机实例。为使实验结果更加精准，其中，令不同规模下产生的实例个数都为 100，并对每一个实例的因子图进行上文所述分割，最终根据算法 2 计算得到该命题公式的结构熵。再计算 100 个实例的结构熵的均值作为该实例组的结构熵。

随着 $\alpha$ 从 0 增至 5，随机产生的 3SAT 实例的结构熵变化情况如图 5 和 6 所示。其中，每个数据点的结构熵均由每个实例组中所有实例的结构熵计算均值得出，每个实例组包含 100 个随机实例。同时，对于每个实例组，统计其中 WP 算法成功收敛的实例个数，得到该实例组的收敛概率。实例收敛的概率随结构熵的变化如图 7 和 8 所示。

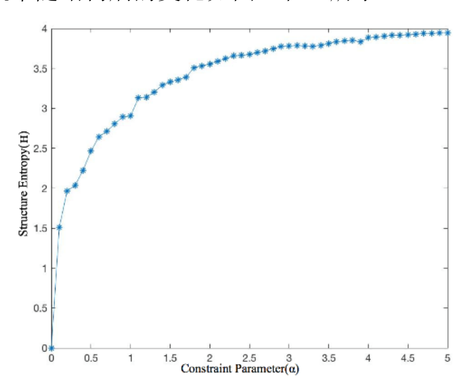



图 5　结构熵分布( $n=10$ )

Fig. 5　Structural entropy distribution( $n=10$ )

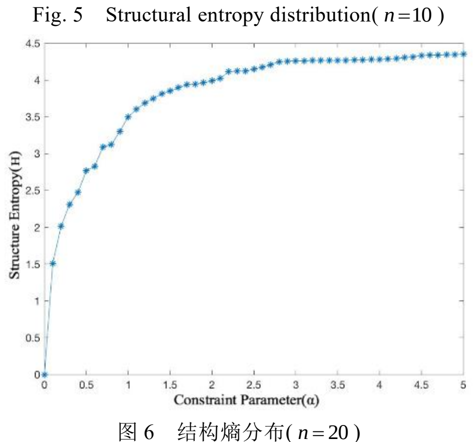



图 6　结构熵分布( $n=20$ )

Fig. 6　Structural entropy distribution( $n=20$ )

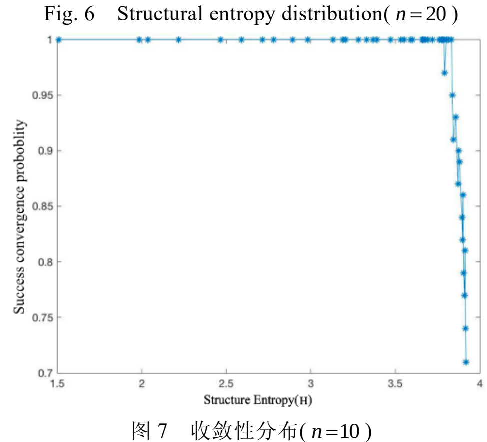



图 7　收敛性分布( $n=10$ )

Fig. 7　Convergent distribution( $n=10$ )

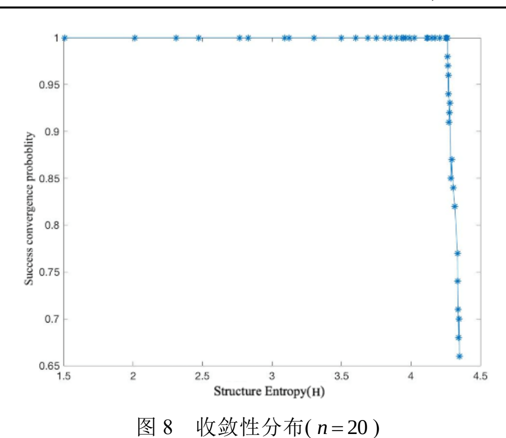



图 8　收敛性分布( $n=20$ )

Fig. 8　Convergent distribution ( $n=20$ )

实验表明，随着 $\alpha$ 增大，随机实例的结构熵也随之增大，但结构熵增量随之减小。由图 5 可知，当 $0<\alpha<2$ 时，结构熵增长较快，当 $\alpha>2$ 时结构熵增幅较为平缓，特别地，当 $\alpha$ 趋近 5 时，结构熵增量趋近于 0。图 7 表明 $\alpha>3$ 时，结构熵变化较为平缓。

当 $n$=10 时，随着 $\alpha$ 趋近于 5，实例的熵值密集分布于 3.5324-3.8941 之间，分布在该区间的实例占实例总数的 74%，最终熵值逼近 3.9。当熵值小于 3.7148 时，WP 算法在所有实例上完全收敛，熵值大于 3.7148 时，算法收敛概率从完全收敛减小到 0.71。

当 $n=20$ 时，实例的熵值密集分布于 3.8136-4.3494 之间，最终逼近 4.35。当熵值小于 4.2795 时，WP 算法几乎完全收敛，当熵值大于 4.2795 时，算法收敛概率急剧下降，最终从 1 减小至 0.66。

结构熵的变化情况表明，当变元规模固定时，随着 $\alpha$ 的增大，因子图的动态复杂性会在一定范围内快速上升，超过阈值时，动态复杂性的上升速度会显著减缓。这是由于在简单因子图结构中，任何的变化，如环路增多，图规模的细微变化，以及因子图直径的改变，都会对因子图结构造成较大影响。而当因子图规模增大到一定程度时，环路的增加以及因子图在一定范围内的规模增大，对于图的动态复杂性影响明显减小，即因子图状态更为稳定。由算法的收敛性变化可知，当因子图状态越为稳定时，WP 算法的收敛性随着因子图的动态复杂度升高而降低。而当因子图的动态复杂性未突破阈值时，WP 算法可以高概率收敛。

因子图的结构决定了 WP 的性能，上述实验给出了不同规模下，WP 算法收敛的充分条件，即当 $n$=10，若 $H^K(G)<3.7148$，WP 算法高概率收敛。当 $n=20$，若 $H^K(G)<4.2795$，WP 算法高概率收敛。基于其他规模因子图的 WP 算法收敛性判定条件，同样可以利用该方法得到。

## 4　结束语

本文提出了 CNF 公式结构信息获取算法，并利用结构熵表达结构信息，通过实验得到了不同因子图规模下结构熵的变化情况以及在不同结构熵范围内 WP 算法在因子图上的收敛性，给出了不同规模下 WP 算法收敛的充分条件。下一步工作将寻找关于命题公式结构熵的临界值或控制参数，用以度量其他信息传播算法的收敛性。并对结构熵和因子图结构信息之间的关系做更进一步的理论分析和实验验证。

## 参考文献：

[1] Martin E. Dyer, Alan M. Frieze, Michael Molloy. A probabilistic analysis of randomly generated binary constraint satisfaction problems [J]. theoretical computer science, 2003, 290 (3): 1815-1828.

[2] Creignou N, Hervé Daudé. The SAT–UNSAT transition for random constraint satisfaction problems [J]. Discrete Mathematics, 2009, 309 (8): 2085-2099.

[3] Martin O C, Monasson R, Zecchina R. Statistical mechanics methods and phase transitions in optimization [J]. Theory Comput Sci, 2001, 265: 3–67.

[4] Cook, Stephen. The Complexity of Theorem-Proving Procedures [C]// In Proceedings of the third annual ACM symposium on Theory of computing, 1971: 151-158.

[5] Clarke E, Biere A, Raimi R, *et al.* Bounded model checking using satisfiability solving [J]. Formal Methods in System Design, 2001, 19 (1): 7-34.

[6] Vizel Y, Weissenbacher G, Malik S. Boolean Satisfiability Solvers and Their Applications in Model Checking [J]. Proceedings of the IEEE, 2015, 103 (11): 2021-2035.

[7] Kuper L, Katz G, Gottschlich J, *et al.* Toward Scalable Verification for Safety-Critical Deep Networks [J]. 2018, 34 (5): 312-322.

[8] Braunstein A, Mezard M, Zecchina R. Survey propagation: an algorithm for satisfiability [J]. Random Structures & Algorithms, 2002, 27 (2): 201-226.

[9] Weiss Y. Correctness of Local Probability Propagation in Graphical Models with Loops [J]. Neural computation, 2000, 12 (1): 1-41.

[10] Weiss Y, Freeman W T. Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology [J]. Neural Computation, 2001, 13 (10): 2173-2200.

[11] 王晓峰，许道云，韦立. 随机可满足实例集上警示传播算法的收敛性 [J]. 软件学报, 2013 (01): 5-15. (Wang Xiaofeng, Xu Daoyun, Wei Li. Randomly satisfy the convergence of the warning propagation algorithm on the instance set [J]. Journal of software, 2013 (01): 5-15)

[12] 王晓峰，许道云. 警示传播算法收敛的充分条件 [J]. 软件学报, 2016 (12): 3003-3013. (Wang Xiaofeng, Xu Daoyun. Sufficient conditions for convergence of warning propagation algorithm [J]. Journal of software, 2016 (12): 3003-3013.)

[13] Su Qinliang, Wu Y C. Convergence analysis of the variance in gaussian belief propagation [J]. IEEE Transactions on Signal Processing, 2014, 62 (19): 5119-5131.

[14] Su Qinliang, Wu Y C. On Convergence Conditions of Gaussian Belief Propagation [J]. IEEE Transactions on Signal Processing, 2015, 63 (5): 1144-1155.

[15] Feige U, Mossel E, Vilenchik D. Complete convergence of message passing algorithms for some satisfiability problems [J]. Theory of computing, 2013, 9 (19): 617-651.

[16] Maneva E, Mossel E, Wainwright M. A new look at survey propagation and its generalizations. [J]. ACM, 2007, 54: 1089–1098.

[17] Shannon C. The lattice theory of information [J]. Transactions of the Ire Professional Group on Information Theory, 2003, 1 (1): 105-107.

[18] Li Angsheng, Pan Yichen. Structural Information and Dynamical Complexity of Networks [J]. IEEE Transactions on Information Theory, 2016, 62 (6): 3290-3339.

[19] Yedidia J S, Freeman W T, Weiss Y. Understanding belief propagation and its generalizations [M]. Morgan Kaufmann, 2003, 8: 239–269.

[20] Kschischang F R, Frey B J, Loeliger H A. Factor graphs and the sum-product algorithm [J]. IEEE Transactions on Information Theory, 2001, 47 (2): 498-519.

[21] 王辛，王晓峰，李卫民. 一种求解最小割的警示传播算法 [J]. 电子学报, 2019, 47 (11): 2386-2391. (Wang Xin, Wang Xiaofeng, Li Weimin. A min-cut algorithm for warning propagation [J]. Acta electronics sinica, 202, 47 (11): 2386-2391.)